SAS Base Exam Certification Practice Exam (Sample)

Study Guide

Everything you need from our exam experts!

Questions

1. How can you sort data in descending order in SAS?

A. By using the SORT procedure with the DESC option

B. By using the SORT procedure with the ASC option

C. By applying the DESC keyword in the DATA step

D. By specifying the ORDER statement

2. What does the KEEP= option do in a SAS procedure?

A. It sorts the variables in ascending order;

B. It specifies which variables to retain in the output dataset;

C. It formats the output as a table;

D. It combines multiple datasets into one;

3. What does the PROC UNIVARIATE procedure analyze?

A. The summary statistics of a variable

B. The distribution of a variable

C. Correlation between multiple variables

D. Data standardization techniques

4. How do you read data from an external file in SAS?

A. By using the INFILE and INPUT statements

B. By importing through the Import Wizard

C. By executing a DATA step using the DATA statement

D. By using the LIBNAME statement to reference files

5. What does the RANK procedure do in SAS?

A. Creates a rank variable based on a specified order

B. Finds the median of numeric variables

C. Summarizes data in frequency tables

D. Sorts data within categories

6. What operator is used to select observations that include a specified substring?

A. IN operator

B. LIKE operator

C. CONTAINS operator

D. SUBSTR operator

7. Which statement correctly describes how PROC SGPANEL interacts with other procedures in SAS?

A. It is independent and does not require other PROC steps

B. It must always be preceded by PROC SORT

C. It can incorporate output from other procedures

D. It can only be used as a standalone procedure

8. Which of the following is NOT a valid type of statement that can be included in a SAS program?

A. Data step

B. PROC step

C. Formula step

D. Global statement

9. What function does the PROC CONTENTS procedure serve in SAS?

A. Outputs the dataset in table format

B. Provides metadata about the dataset

C. Runs a statistical analysis on the dataset

D. Graphs the dataset distribution

10. What does the term 'operand' in an expression represent?

A. Constants or variables

B. A sequence of numbers

C. Logical statements

D. Data types

Answers

1. A
2. B
3. B
4. A
5. A
6. C
7. C
8. C
9. B
10. A

Explanations

1. How can you sort data in descending order in SAS?

A. By using the SORT procedure with the DESC option

B. By using the SORT procedure with the ASC option

C. By applying the DESC keyword in the DATA step

D. By specifying the ORDER statement

To sort data in descending order in SAS, the SORT procedure is utilized with the DESC option. This procedure is specifically designed for sorting datasets, allowing for customization in the order of the results. When you invoke the SORT procedure, you can specify the variable(s) by which you wish to sort your data, and by adding the DESC option after the variable name, SAS understands that you want the sort to be performed in descending order. The options in the question indicate different methods, but only using the SORT procedure with the DESC option correctly implements sorting in descending order according to SAS syntax. Using the ASC option, for instance, would initiate sorting in ascending order, while the DATA step does not directly support a descending sort without the use of a separate SORT procedure. The ORDER statement is also not a valid method for achieving this within a SORT context in SAS. Thus, the correct approach is to utilize the SORT procedure with the DESC option specifically for this purpose.

2. What does the KEEP= option do in a SAS procedure?

A. It sorts the variables in ascending order;

B. It specifies which variables to retain in the output dataset;

C. It formats the output as a table;

D. It combines multiple datasets into one;

The KEEP= option in a SAS procedure is specifically designed to dictate which variables are to be retained in the output dataset. When you use this option, you can list the variables you want to preserve, and all other variables will be excluded from the output. This can be particularly useful when working with large datasets, as it allows you to focus only on the relevant variables you need for your analysis or reporting without altering the original dataset. For instance, if your dataset contains many variables but your analysis requires only a subset of those, using the KEEP= option helps streamline your output, making it more manageable and easier to work with. This feature enhances the efficiency of the data management process in SAS and ensures clarity by including only necessary information in your output datasets. The other options provided do not accurately describe the function of the KEEP= option. Some pertain to different SAS functionalities, such as sorting, formatting, or dataset combination, but they do not relate to the retention of specific variables in the output, which is the core purpose of the KEEP= option.

3. What does the PROC UNIVARIATE procedure analyze?

A. The summary statistics of a variable

B. The distribution of a variable

C. Correlation between multiple variables

D. Data standardization techniques

The PROC UNIVARIATE procedure in SAS is designed specifically to analyze the distribution of a variable. It provides a comprehensive overview of the variable's distribution characteristics through various statistical measures, such as mean, median, mode, standard deviation, and percentiles. Additionally, PROC UNIVARIATE produces visual outputs like histograms, box plots, and quantile-quantile plots, which help in understanding the shape and variability of the data distribution. While summary statistics of a variable are indeed part of what PROC UNIVARIATE provides, the primary focus is on the distribution itself—how the data are spread and their underlying patterns. Correlation between multiple variables is not within the scope of PROC UNIVARIATE; this analysis is more suitable for procedures like PROC CORR. Similarly, data standardization techniques are not covered by this procedure; they are typically handled in other methods, such as transformations and scaling that are outside of PROC UNIVARIATE's intended functionality.

4. How do you read data from an external file in SAS?

A. By using the INFILE and INPUT statements

B. By importing through the Import Wizard

C. By executing a DATA step using the DATA statement

D. By using the LIBNAME statement to reference files

The correct method to read data from an external file in SAS involves using the INFILE and INPUT statements. The INFILE statement specifies the external file's location, while the INPUT statement defines how the data should be read and formatted into SAS variables. This approach allows for precise control over how data is processed, including handling different delimiters, reading specific records, and managing formatted data. In contrast, while importing through the Import Wizard (the second choice) can also read data, this method is more of a graphical interface approach suited for those who prefer not to write code manually. It may not provide the same level of scriptable flexibility as using INFILE and INPUT. Executing a DATA step using the DATA statement (the third choice) is relevant but is not the complete answer. While a DATA step is essential to create SAS datasets, it needs the INFILE and INPUT statements to actually read data from an external source. The LIBNAME statement (the fourth choice) is used to assign a library reference to a directory containing data files but does not directly read or import data itself. It setup references rather than define the data import process. Therefore, the combination of the INFILE and INPUT statements in a DATA step is the correct approach for reading external data

5. What does the RANK procedure do in SAS?

A. Creates a rank variable based on a specified order

B. Finds the median of numeric variables

C. Summarizes data in frequency tables

D. Sorts data within categories

The RANK procedure in SAS is primarily designed to create a rank variable based on a specified order. When you use this procedure, you can assign ranks to the values in a data set according to their magnitude, either in ascending or descending order. This is particularly useful in scenarios where you need to order data for analysis or comparison, such as determining the position of items or individuals relative to each other based on a specific variable. This procedure allows for various ranking methods, including handling ties. With the RANK procedure, you can customize the way ranks are assigned so that they are meaningful for your specific analytical needs. As a result, this option is not only fundamental to the functionality of the RANK procedure, but it is also widely applicable in statistical analyses where ranking of observations plays a critical role. Other options mention operations like finding medians, summarizing data, and sorting, which are functionalities associated with other procedures or functions in SAS, but they do not accurately describe the purpose of the RANK procedure. The RANK procedure is specifically designed for generating rank variables, making it the clear choice in this context.

6. What operator is used to select observations that include a specified substring?

A. IN operator

B. LIKE operator

C. CONTAINS operator

D. SUBSTR operator

The operator used to select observations that include a specified substring is the CONTAINS operator. This operator is particularly useful in situations where you want to find records within a character variable that contain a certain subset of characters, regardless of their position within the variable. When utilizing the CONTAINS operator, you can effectively filter data to find any instance where a specific substring appears within a larger string. This capability is essential for text analysis or when handling data with varying descriptions that may meet certain criteria. Although other operators like the LIKE operator may seem relevant, the LIKE operator is primarily used in conjunction with pattern matching in SQL queries, often varying from the straightforward substring search that the CONTAINS operator provides. The IN operator is generally used for matching values in lists, and the SUBSTR operator is designed for extracting substrings rather than selecting observations based on the presence of a substring. This specificity of the CONTAINS operator makes it the most suitable choice for the question asked.

7. Which statement correctly describes how PROC SGPANEL interacts with other procedures in SAS?

A. It is independent and does not require other PROC steps

B. It must always be preceded by PROC SORT

C. It can incorporate output from other procedures

D. It can only be used as a standalone procedure

PROC SGPANEL is designed to facilitate the creation of panel data graphics in SAS, allowing users to visualize data easily by displaying different variables or groups in separate panels. One of the defining features of PROC SGPANEL is its ability to incorporate output from other procedures, specifically through the use of data sets that they produce. For instance, you can use PROC SUMMARY or PROC MEANS to generate summary statistics and then feed that output directly into PROC SGPANEL for graphical representation. This integration enhances the analytic workflow by allowing for the combination of statistical analysis and visual representation in a cohesive manner. Unlike other options, which imply limitations on the procedure's capabilities, PROC SGPANEL's interaction with other PROC steps signifies its versatility and ability to enhance the overall data analysis process within SAS.

8. Which of the following is NOT a valid type of statement that can be included in a SAS program?

A. Data step

B. PROC step

C. Formula step

D. Global statement

In SAS programming, there are specific types of statements that dictate how data is processed and managed within a program. The data step is critical for creating, manipulating, and transforming datasets. Similarly, a PROC step is essential for analyzing data, producing reports, and utilizing SAS procedures, like those for statistical analysis. Global statements are also valid in a SAS program, serving to define global options or settings that apply to subsequent data or PROC steps throughout the session, such as the OPTIONS statement. The term "Formula step," however, is not recognized as a valid statement type within SAS. While SAS does allow for the creation of formulas in various contexts (such as in DATA steps or PROC steps for calculations), there isn't a standalone statement classified as a "Formula step." This distinction makes it clear that "Formula step" does not align with the established terminology and functionality within SAS, confirming it as the correct choice.

9. What function does the PROC CONTENTS procedure serve in SAS?

A. Outputs the dataset in table format

B. Provides metadata about the dataset

C. Runs a statistical analysis on the dataset

D. Graphs the dataset distribution

The PROC CONTENTS procedure in SAS is specifically designed to provide metadata about a dataset. This includes essential information such as the dataset's name, the number of observations and variables, variable attributes (like types and lengths), and the indexing information. This metadata is crucial for understanding the structure and characteristics of the dataset, allowing users to make informed decisions about data analysis and management. For those exploring data within SAS, having access to metadata through PROC CONTENTS helps facilitate data exploration, quality checks, and informed data manipulation. Understanding the contents of a dataset allows users to tailor their analysis effectively and ensure they are working with the correct data structures. The other options do not accurately represent the function of PROC CONTENTS. While one might output the dataset or conduct statistical analyses, those tasks are handled by different procedures in SAS, such as PROC PRINT for displaying data or PROC MEANS for statistical analysis. The graphing of distributions is typically done through PROC SGPLOT or similar procedures, focusing on visual representation rather than the metadata aspect that PROC CONTENTS addresses.

10. What does the term 'operand' in an expression represent?

A. Constants or variables

B. A sequence of numbers

C. Logical statements

D. Data types

The term 'operand' in an expression refers to the components that the operators manipulate to produce a result. This includes both constants and variables, which are the building blocks of expressions used in programming and mathematical computations. For example, in the expression `5 + x`, the number `5` is a constant operand, while `x` can be a variable operand. Both of these types of operands can be involved in operations for calculations. The flexibility of operands allows for expressions to be dynamic, incorporating variable values that can change as the program executes. Other options, while relevant to programming and mathematical contexts, do not capture the full definition of operands. A sequence of numbers specifically does not encompass variables, as operands can be both numbers and variable names. Logical statements represent conditions that evaluate to true or false but are not classified as operands themselves. Data types refer to the kind of data a variable can hold, like integers or characters, rather than the actual values that serve as operands in an operation.